

Compositions and Methods for Protein Isolation

FIELD OF THE INVENTION

5 The invention relates in general to improved methods of protein isolation and identification of protein binding partners for a protein of interest.

BACKGROUND OF THE INVENTION

10 Identification of protein/protein interactions is at the core of understanding the biological processes occurring in living cells. Traditionally, the potential interacting proteins have been identified by genetic methods (two hybrid screens) with subsequent verification of the interaction by co-immunoprecipitation. While this method has been very successful for detection of two interacting proteins, it is of limited utility when more complex protein aggregates such as ribosomes, splice complexes or transcription complexes are investigated.

15 To identify and isolate yeast complex protein aggregates, an alternative method has been developed by Seraphin et al. (Rigaut et al., 1999, Nature Biotech., 17:1030-1032; Puig et al., 2001, Methods, 24: 218-219; U.S. 2002/0061513, reviewed in Terpe et al., 2003, App. Microbiol. Biotechnol., 60:523-533). This method combines purification of the protein complex of interest using two different affinity purification tags fused to at least one known protein
20 component of a complex of interest by genetic methods, with subsequent mass spectroscopy to identify the unknown components of the isolated complex. The use of two consecutive purification steps allows for isolation of the complex, in a purified form, without disruption of the targeted complex. Only certain combinations of purification tags are suitable for this method.

25 The calmodulin-binding domain of the calmodulin binding peptide (CBP-tag) and the IgG binding domain(s) of Staphylococcus aureus protein A represent an efficient combination of purification tags, according to this method (Rigaut et al., supra; Puig et al., supra; U.S. 2002/0061513). The interaction between the CBP-tag and the purification matrix (immobilized calmodulin) can be controlled by the presence of Ca^{2+} . In the presence of Ca^{2+} , the CBP tag

binds to the purification matrix whereas removal of Ca^{2+} with a chelating agent such as EGTA, allows recovery of the tagged protein from the purification resin under mild conditions (Stofko-Hahn et al., 1992, FEBS Lett., 302:274-278). The IgG-binding domain of protein A provides specific, high affinity binding with little non-specific interaction. However, it is very difficult to elute protein A tagged proteins from IgG-columns. Consequently, elution can only be achieved by removing protein A fusion proteins by digestion with a site-specific protease. Utilization of the IgG-binding domain of protein A therefore requires additional processing steps and leads to contamination of the purified protein with the protease.

There is a need in the art for a method to detect and identify protein complexes that does not disrupt protein-protein interactions. This method will also facilitate detection of binding partners for a protein of interest in the absence of prior knowledge of the binding partner(s) or the function of the protein complex. There is also a need in the art for a purification protocol for protein complexes that does not require digestion with a protease enzyme. This method provides a simple, generic purification protocol that can be used routinely, and, possibly, in an automated system, for the purification of protein complexes and for proteome analysis.

SUMMARY OF THE INVENTION

The invention provides reagents for detecting and isolating proteins in a complex. In particular, the invention provides for a vector comprising at least two affinity tags. The invention provides for a protein comprising at least two affinity tags. Alternatively, the invention provides for a protein of interest comprising at least one affinity tag, and a binding partner, or candidate binding partner for the protein of interest comprising at least a second affinity tag. The invention also provides methods for identifying and detecting a protein in a complex, without disruption of the complex. The method of the invention can be used to find one or more "target" binding partners for a "bait" protein of interest. According to the method of the invention, the protein of interest is fused in frame, either N-terminally, C-terminally or a combination thereof, to at least two affinity tags.

In one embodiment, the invention provides for a polynucleotide comprising at least two affinity tag sequences. In one embodiment, one of the tag sequences encodes streptavidin-

binding peptide having a nucleotide sequence presented in Figure 1. The at least two tag sequences are either directly adjacent to each other or are separated by a spacer, for example, of 1-60 nucleotides. Either of the first or second tags can be located 5' of the other tag.

In one embodiment the invention provides for a polynucleotide comprising a gene of interest and at least two tag sequences. The gene of interest is fused in frame with each of the tag sequences. In one embodiment, one of the tag sequences encodes streptavidin-binding peptide having a nucleotide sequence presented in Figure 1.

As used herein, the terms "nucleic acid", "polynucleotide" and "oligonucleotide" refer to polydeoxyribonucleotides (containing 2-deoxy-D-ribose), to polyribonucleotides (containing D-ribose), and to any other type of polynucleotide which is an N-glycoside of a purine or pyrimidine base, or modified purine or pyrimidine bases (including abasic sites). There is no intended distinction in length between the term "nucleic acid", "polynucleotide" and "oligonucleotide", and these terms will be used interchangeably. These terms refer only to the primary structure of the molecule. Thus, these terms include double- and single-stranded DNA, as well as double- and single-stranded RNA.

As used herein, "protein of interest" means any protein for which the nucleic acid sequence is known or available, or that becomes available, such that it can be cloned into a nucleic acid vector which is suitable for expression in the appropriate host cells or cell-free expression systems. For purification of a protein complex, the nucleic acid sequence of at least one of the subunits of the protein complex must be known or available.

The invention also provides for identification and/or purification of a protein complex, or identification and/or purification of a complex of one or more proteins and one or more biomolecules. As used herein, a "biomolecule" includes a protein, peptide, nucleic acid, antibody, or other biomolecule. A biomolecule complex is a complex of at least two biomolecules, preferably at least one protein in association with either other proteins or with other biomolecules, for example, nucleic acid or antibody. The biomolecule complexes can be naturally occurring, such as nuclear snRNPs or antigen-antibody complexes, or they can be non-naturally occurring, for example, mutant DNA binding protein in association with mutant target DNA. Any complex molecule comprising as one or more subunits a polypeptide or subunit expressed according to the invention and/or further comprising other components which

associate in a manner stable enough to remain associated during the affinity purification steps is a biomolecule complex that can be detected/purified by the method of the invention.

The terms “tag” or “affinity tag” are used interchangeably herein. As used herein, “tag” or “affinity tag” means a moiety that is fused in frame to the 5' or 3' end of, or internally to, the protein product of a gene of interest, a biomolecule of the invention, or another tag. A “tag” specifically binds to a ligand as a result of attractive forces that exist between the tag and a ligand. “Specifically binds” as it refers to a “tag” and a ligand means via covalent or hydrogen bonding or electrostatic attraction or via an interaction between for example a tag and a ligand, an antibody and an antigen, protein subunits, or a nucleic acid binding protein and a nucleic acid binding site. Preferably, a “tag” of the invention, binds a ligand with a dissociation constant (K_D) of at least about $1 \times 10^3 \text{ M}^{-1}$, usually at least $1 \times 10^4 \text{ M}^{-1}$, typically at least $1 \times 10^5 \text{ M}^{-1}$, preferably at least $1 \times 10^6 \text{ M}^{-1}$ to $1 \times 10^9 \text{ M}^{-1}$ or more, for example $1 \times 10^{14} \text{ M}^{-1}$ for streptavidin-avidin binding, $1 \times 10^{15} \text{ M}^{-1}$, $1 \times 10^{16} \text{ M}^{-1}$, $1 \times 10^{20} \text{ M}^{-1}$, or more. A tag does not interfere with expression, folding or processing of the tagged protein or with the ability of a protein to bind to its binding partner. Tags include but are not limited to calmodulin binding peptide, streptavidin binding peptide, calmodulin binding peptide, streptavidin, avidin, polyhistidine tag, polyarginine tag, FLAG tag, c-myc tag, S-tag, cellulose binding domain, chitin-binding domain, glutathione S-transferase tag, maltose-binding protein, TrxA, DsbA, hemagglutinin epitope, InaD, NorpA, and GFP (see Honey et al., supra; Hu et al., supra; Puig et al., supra; Rigaut et al., supra; Terpe, supra; U.S. 2002/0061513, Kimple et al., Biotechniques. 2002, 33:578) incorporated by reference herein in their entirety.

As used herein, “fused in frame” means fused such that the correct translational reading frame is maintained thereby allowing for expression of all of the components of the chimeric or fusion protein.

As used herein, the term “fused to the amino-terminal end” refers to the linkage of a polypeptide sequence to the amino terminus of another polypeptide. The linkage may be direct or may be mediated by a short (e.g., about 2-20 amino acids) linker peptide. Examples of useful linker peptides include, but are not limited to, glycine polymers $((G)_n)$ including glycine-serine and glycine-alanine polymers. It should be understood that the amino-terminal end as used herein refers to the existing amino-terminal amino acid of a polypeptide, whether or not that

amino acid is the amino terminal amino acid of the wild type or a variant form (e.g., an amino-terminal truncated form) of a given polypeptide.

As used herein, the term “fused to the carboxy-terminal end” refers to the linkage of a polypeptide sequence to the carboxyl terminus of another polypeptide. The linkage may be direct or may be mediated by a linker peptide. As with fusion to the amino-terminal end, fusion to the carboxy-terminal end refers to linkage to the existing carboxy-terminal amino acid of a polypeptide.

As used herein, streptavidin binding peptide (SBP)” or streptavidin binding protein means a synthetic streptavidin-binding domain that binds streptavidin with a dissociation constant from $1 \times 10^5 \text{ M}^{-1}$ - $5 \times 10^{10} \text{ M}^{-1}$ (for example, $1 \times 10^5 \text{ M}^{-1}$, $1 \times 10^6 \text{ M}^{-1}$, $1 \times 10^7 \text{ M}^{-1}$, $1 \times 10^8 \text{ M}^{-1}$, $1 \times 10^9 \text{ M}^{-1}$, $1 \times 10^{10} \text{ M}^{-1}$ in the absence but not in the presence of biotin. In one embodiment, SBP has the amino acid sequence presented in Figure 1. Additional SBP sequences useful according to the invention include SB1, SB2, SB5, SB9, SB11 and SB12 (Wilson et al., 2001, Proc. Natl. Acad. Sci USA, 98:3750), presented in Figure 2.

The invention also provides for an isolated polynucleotide comprising at least two tag sequences, wherein one of the tag sequences encodes streptavidin binding peptide and the other encodes calmodulin binding peptide. The at least two tag sequences are either directly adjacent to each other or are separated by a spacer, for example, of 1-60 nucleotides. Either of the streptavidin binding peptide tag or the calmodulin binding peptide tag can be located 5' of the other tag.

The invention also provides for an isolated polynucleotide comprising a gene sequence of interest and at least two tag sequences fused in frame with each other. One of the two tag sequences encodes streptavidin binding peptide and one of the tag sequences encodes calmodulin binding peptide.

As used herein, “calmodulin binding peptide (CBP)” or calmodulin binding peptide means a peptide that binds calmodulin, preferably with a dissociation constant from $1 \times 10^3 \text{ M}^{-1}$ to $1 \times 10^{14} \text{ M}^{-1}$ and preferably $1 \times 10^6 \text{ M}^{-1}$ to $1 \times 10^{10} \text{ M}^{-1}$ and more preferably, $1 \times 10^7 \text{ M}^{-1}$ to $1 \times 10^9 \text{ M}^{-1}$, in a Ca^{2+} dependent manner. Binding occurs in the presence of Ca^{2+} , in the range of 0.1 μM to 10mM. CBP is derived from the C-terminus of skeletal-muscle myosin light chain kinase. In the

presence of Ca^{2+} , the CBP tag binds to calmodulin and, upon removal of Ca^{2+} , for example, in the presence of a chelating agent such as EGTA (preferably in the range of 0.1 μM to 10mM), CBP does not bind calmodulin. In one embodiment, CBP has the amino acid sequence presented in Figure 1. Additional CBP sequences useful according to the invention include: bovine
5 neuromodulin AA 37-53 KIQASFRGHITRKKLKG (Hinfichsen et al., 1993, Proc. Natl. Acad Sci USA, 90:1585); calmodulin-dependent protein kinase I (CMKI) AA 294-318
SEIQKKNFAKSKWKQAFNATAVVRHMRK; calmodulin-dependent protein kinase II
(CMKII) AA 290-309 LKKFNARRKLKGAILTTMLA; and tuberous sclerosis 2 (TSC)
WIARLRHIKRLRQRIL (Noonan et al., 2002, Arch. Biochem. Biophys. 389:32).

10 In one embodiment, each of the tags of the isolated polynucleotide are adjacent to the 5' end of the target gene sequence.

In another embodiment, each of the tags of the isolated polynucleotide are adjacent to the 3' end of the target gene sequence.

15 Since mononucleotides are reacted to make oligonucleotides in a manner such that the 5' phosphate of one mononucleotide pentose ring is attached to the 3' oxygen of its neighbor in one direction via a phosphodiester linkage, an end of an oligonucleotide is referred to as the "5' end" if its 5' phosphate is not linked to the 3' oxygen of a mononucleotide pentose rings, and as the "3' end" if its 3' oxygen is not linked to a 5' phosphate of a subsequent mononucleotide pentose ring.

20 As used herein, "adjacent" or "tandem" means immediately preceding or following. "Adjacent" also means preceding or following and separated by a linker, for example a nucleic acid linker of 6-60 nucleic acids or an amino acid linker of 2-20 amino acids.

The invention also provides for a vector comprising the isolated polynucleotides of the invention.

25 As used herein, "vector" means a cloning vector that contains the necessary regulatory sequences to allow transcription and translation of a cloned gene or genes.

The invention also provides for a cell comprising the vector of the invention.

The invention also provides for a composition comprising the isolated polynucleotides of the invention.

The invention also provides for a chimeric protein comprising at least two affinity tags, wherein one of the tags is streptavidin binding peptide having the sequence presented in Figure 1. The at least two tags are either directly adjacent to each other or are separated by a spacer, as defined herein. Either of the first or second tags can be located N-terminal to the other tag.

The invention also provides for a chimeric protein comprising a protein of interest fused in frame to at least two different affinity tags, one of which is streptavidin binding peptide having the sequence presented in Figure 1.

The invention also provides for a chimeric protein comprising a streptavidin binding peptide and a calmodulin binding peptide. The tags are either directly adjacent to each other or are separated by a spacer, as defined herein. Either of the first or second tags can be located N-terminal to the other tag.

The invention also provides for a chimeric protein comprising a protein of interest fused in frame to at least two different affinity tags, one of which is streptavidin binding peptide, and wherein one of the affinity tags is calmodulin binding peptide.

In one embodiment, each of the tags are adjacent to the N-terminus of the protein of interest.

In another embodiment, each of the tags are adjacent to the C-terminus of the protein of interest.

As used herein, a “chimera” or “fusion” means a fusion of a first amino acid sequence (protein) comprising a protein product of a gene of interest, joined to a second amino acid sequence encoding a first tag, and joined to at least a third amino acid sequence encoding a second tag. A “chimera” according to the invention contains three or more amino acid sequences (for example a sequence encoding a protein of interest, a sequence encoding calmodulin-binding peptide and a sequence encoding streptavidin-binding peptide) from unrelated proteins, joined to form a new functional protein. A chimera of the invention may present a foreign polypeptide which is found (albeit in a different protein) in an organism which also expresses the first

protein, or it may be an "interspecies", "intergenic", etc. fusion of protein structures expressed by different kinds of organisms. The invention encompasses chimeras wherein at least two tag amino acid sequences are joined N-terminally or C-terminally to the protein product of the gene of interest, or wherein a first tag sequence is joined N-terminally and a second tag sequence is joined C-terminally to a protein product of a gene of interest. A "chimera" of the invention includes a protein of interest fused to at least two tags, wherein the tags are located N- or C-terminally, or any combination thereof. The invention also encompasses a chimera wherein one or more of the tag amino acid sequences are fused internally to the amino acid sequence of a protein of interest.

A "chimera" according to the invention also refers to a fusion of a first amino acid sequence comprising a protein product of a gene of interest, joined to at least a second amino acid sequence encoding at least one tag of the invention.

As used herein, "chimeric or fusion protein or polypeptide" refers to a heterologous amino acid sequence of two or more "tag" amino acid sequences fused in frame to the amino acid sequence of interest. In one embodiment, the two or more tag amino acid sequences are fused to the N or C termini of the amino acid sequence of the protein of interest. In one embodiment, a first tag amino acid sequence is fused in frame to the N-terminus of the amino acid sequence of the protein of interest and the second tag amino acid sequence is fused in frame to the C-terminus of the protein of interest. The invention also provides for a first chimeric protein comprising a first tag amino acid sequence fused to a first protein of a complex and a second chimeric protein comprising a second tag amino acid sequence fused to a second protein, wherein the first and second protein are present in the same complex.

The invention also provides for a composition comprising the isolated chimeric proteins of the invention.

The invention also provides for a method of detecting or isolating one or more binding partners for a protein encoded by a gene of interest, comprising the following steps. A gene sequence of interest is cloned into a vector such that the gene of interest is fused in frame with at least two different tag sequences. One of the tag sequences encodes streptavidin binding peptide having the amino acid sequence presented in Figure 1. The vector is introduced into a cell

comprising at least one candidate binding partner. The protein product of the gene of interest and the candidate binding partner are allowed to form a complex in the cell. The complex is isolated by lysing the cells and performing at least one round of affinity purification. The protein complex is then detected.

5 The invention also provides for a method of detecting or isolating one or more binding partners for a protein encoded by a gene of interest, comprising the following steps. A gene sequence of interest is cloned into a vector such that the gene of interest is fused in frame with at least two different tag sequences. One of the tag sequences encodes streptavidin binding peptide and one of the tag sequences encodes calmodulin-binding peptide. The vector is introduced into
10 a cell comprising at least one candidate binding partner. The protein product of the gene of interest and the candidate binding partner are allowed to form a complex in the cell. The complex is isolated by lysing the cells and performing at least one round of affinity purification. The protein complex is then detected.

 In one embodiment, the cell comprises a vector that expresses at least one candidate
15 binding partner for the protein product of the gene of interest.

 In one embodiment the candidate binding partner expresses a tag.

 The invention also provides for a method of detecting or isolating a protein complex comprising the following steps. A gene sequence of interest is cloned into a vector such that the gene sequence of interest is fused in frame with at least two different tag sequences. One of the
20 two tag sequences encodes streptavidin binding peptide having the amino acid sequence presented in Figure 1. The vector is introduced into a cell that expresses at least one protein binding partner for the protein product of the gene sequence of interest. The protein product of the gene of interest and the protein binding partner are allowed to form a complex. The complex is isolated by lysing the cells and performing at least one round of affinity purification.

25 The invention also provides for a method of detecting or isolating a protein complex comprising the following steps. A gene sequence of interest is cloned into a vector such that the gene sequence of interest is fused in frame with at least two different tag sequences. One of the two tag sequences encodes streptavidin binding peptide and one of the two tag sequences

encodes calmodulin binding peptide. The vector is introduced into a cell that expresses at least one protein binding partner for the protein product of the gene sequence of interest. The protein product of the gene of interest and the protein binding partner are allowed to form a complex. The complex is isolated by lysing the cells and performing at least one round of affinity purification.

In one embodiment, the cell comprises a vector that expresses at least one candidate binding partner for the protein product of the gene of interest.

In one embodiment, the candidate binding partner comprises a tag.

In another embodiment, the complex is isolated by performing at least two successive rounds of affinity purification.

As used herein, “protein complex” means two or more proteins or biomolecules that are associated. As used herein, “associated” as it refers to binding of two or more proteins or biomolecules, means specifically bound by hydrogen bonding, covalent bonding, or via an interaction between, for example a protein and a ligand, an antibody and an antigen, protein subunits, or nucleic acid and protein. Under conditions of stable association, binding results in the formation of a protein complex, under suitable conditions, with a dissociation constant, (K_D) of at least about $1 \times 10^3 \text{ M}^{-1}$, usually at least $1 \times 10^4 \text{ M}^{-1}$, typically at least $1 \times 10^5 \text{ M}^{-1}$, preferably at least $1 \times 10^6 \text{ M}^{-1}$ to $1 \times 10^7 \text{ M}^{-1}$ or more, for example $1 \times 10^{14} \text{ M}^{-1}$, $1 \times 10^{16} \text{ M}^{-1}$, $1 \times 10^{18} \text{ M}^{-1}$, $1 \times 10^{20} \text{ M}^{-1}$, $1 \times 10^{30} \text{ M}^{-1}$ or more, for each member of the complex. Methods of performing binding reactions between members of a protein complex, as defined herein, are well-known in the art and are described hereinbelow.

As used herein, “form a complex” means to incubate members of a protein complex under conditions, for example, in the presence of the appropriate buffer, salt conditions, and pH, that allow for association of the members of the protein complex. “Form a complex” also means to bind, under suitable conditions, with a dissociation constant (K_D) of at least about $1 \times 10^3 \text{ M}^{-1}$, usually at least $1 \times 10^4 \text{ M}^{-1}$, typically at least $1 \times 10^5 \text{ M}^{-1}$, preferably at least $1 \times 10^6 \text{ M}^{-1}$ to $1 \times 10^7 \text{ M}^{-1}$, for example $1 \times 10^{14} \text{ M}^{-1}$, $1 \times 10^{16} \text{ M}^{-1}$, $1 \times 10^{18} \text{ M}^{-1}$, $1 \times 10^{20} \text{ M}^{-1}$, $1 \times 10^{30} \text{ M}^{-1}$ or more, for each member of the complex.

As used herein, “affinity purification” means purification of a complex via binding of at least one of the affinity tags of a member of the complex to the ligand for the affinity tag. In one embodiment, the tag is associated with a support material. In a preferred embodiment, the method of the invention utilizes at least two affinity purification steps.

5 As used herein, “purification resin” or “affinity purification resin” refers to a support material to which a ligand of the invention is immobilized. A “purification resin” according to the invention includes but is not limited to beaded derivatives of agarose, cellulose, polystyrene gels, cross-linked dextrans, polyacrylamide gels, and porous silica.

Further features and advantages of the invention are as follows.

10 BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the sequence of the CBP/SBP tandem affinity tags.

Figure 2 is a Table presenting SBP sequences useful according to the invention.

Figure 3(a) and 3(b) show expression vectors comprising nucleic acids encoding CBP and SBP
15 affinity tags useful according to the invention.

Figure 4(a) and 4(b) show expression vectors for expression of a “target” binding partner of the invention.

Figure 5 is a Western blot of affinity purified Mef2c-FLAG.

Figure 6 is a Tris-glycine acrylamide gel of affinity purified Mef2A/Mef2c.

20 DESCRIPTION

The invention provides for a method of detecting and/or purifying a protein complex under mild conditions such that the complex is not dissociated. The purification methods described herein allow for isolation of a protein complex that maintains functional activity. The
25 invention also provides for detection of binding partners for a protein of interest.

Tags

The invention provides an affinity purification tag system comprising an SBP-tag having an amino acid sequence presented in Figure 1. A second affinity tag includes but is not limited to any of the tags described herein. The invention also provides an affinity purification tag system combining a CBP-tag with an SBP-tag. The invention also provides for an SBP having a sequence presented in any of Luo et al., 1998, J. Biotechnol., 65:225-228; Devlin et al., 1990, Science, 249:404-406; Ostergaard et al., 1995, FEBS Lett, 362:306-308; Gissel et al., 1995, J Pept Sci., 1:217-226; Schmidt et al., 1996, J Mo Biol., 255:753-766; Skerra et al., 1996, Biomol Eng., 16:79-86; Koo et al., 1998, Appl Environ Microbiol., 64:2490-2496; Aubrey et al., 2001, Biol Chem., 383:1621-1628. Preferably, the invention provides for an affinity purification tag system comprising an SBP tag and at least a second affinity tag. Other SBP tags useful according to the invention are presented in Figure 2, in particular SB1, SB2, SB5, SB9, SB11 and SB12.

Streptavidin has traditionally been used as an affinity tag because it binds biotin with high affinity ($K_d = 10^{-14}$ M) and specificity. Streptavidin will bind biotinylated compounds (such as proteins and nucleic acids) under physiological conditions and the bound compounds are subsequently eluted with biotin. Tagging the targeted protein for streptavidin purification can be achieved by several methods. Biotinylation can be directed to the tagged protein by using domains that are substrates for biotin ligases (de Boer et al., 2003, Proc Natl Acad Sci USA, 100:7480-7485)). However, this approach requires a biotin ligase, which has to be delivered either *in vivo* or *in vitro* (de Boer et al., *supra*). Alternatively, protein tags can be used that have affinity for streptavidin in the absence but not in the presence of biotin and are thus elutable. Two tags with such features have been described: streptag II (Schmidt et al., 1996, J Mol Biol., 225:753-766) and the streptavidin binding peptide (SBP) (Wilson et al., 2001, Proc Natl Acad Sci USA, 98:3750-3755; Keefe et al., 2001, Protein Expr Purif., 23:440-446; U.S. 2002/0155578 A1)). SBP has a much higher affinity for streptavidin than streptag II (Wilson et al., *supra*).

CBP has 26 residues (see Figure 1) and is derived from the C-terminus of skeletal-muscle myosin light chain kinase, which binds calmodulin with nanomolar affinity in the presence of 0.2mM CaCl₂ (Blumenthal et al., Proc. Natl. Acad Sci USA, 82:3187-3191). In one embodiment of the invention, CBP has the sequence presented in Figure 1. Additional CBP sequences useful

according to the invention include: bovine neuromodulin AA 37-53 KIQASFRGHITRKKLKG (Hinfichsen et al., 1993, Proc. Natl. Acad Sci USA, 90:1585); calmodulin-dependent protein kinase I (CMKI) AA 294-318 SEQIKKNFAKSKWKQAFNATAVVRHMRK; calmodulin-dependent protein kinase II (CMKII) AA 290-309 LKKFNARRKLKGAILTTMLA; and
5 tuberous sclerosis 2 (TSC) WIARLRHIKRLRQRIL (Noonan et al., 2002, Arch, Biochem. Biophys. 389:32).

A purification tag, according to the invention, possesses the following characteristics: (i) the interaction between the tag and the purification matrix is high affinity for example, in the range of 10^3M^{-1} to 10^{14}M^{-1} ; or more (ii) binding occurs under physiological conditions, and does
10 not disrupt the protein–protein interactions of the targeted complex; (iii) elution of the targeted complex from the purification matrix occurs under physiological conditions that do not disrupt the protein–protein interactions; (iv) the binding and elution conditions of the two purification tags are compatible with each other; and (v) the purification tag and the purification matrix have low affinity, for example, less than 10^3M^{-1} , for other proteins within the cell lysate to reduce
15 non-specific background.

The invention provides for fusion proteins that are tagged with at least two adjacent tag moieties. In a preferred embodiment, a protein of interest is tagged at the N- or C-terminus with adjacent SBP and CBP tags. Combinations of any of the following tags are also useful according to the invention: calmodulin binding peptide, streptavidin binding peptide, calmodulin binding
20 peptide, streptavidin, avidin, polyhistidine tag, polyarginine tag, FLAG tag, c-myc tag, S-tag, cellulose binding domain, chitin-binding domain, glutathione S-transferase tag, Maltose-binding protein, TrxA, DsbA, hemagglutinin epitope, InaD, NorpA, and GFP.

The invention also provides for a first protein that is tagged with at least one of the following tags: calmodulin binding peptide, streptavidin binding peptide, calmodulin binding
25 peptide, streptavidin, avidin, polyhistidine tag, polyarginine tag, FLAG tag, c-myc tag, S-tag, cellulose binding domain, chitin-binding domain, glutathione S-transferase tag, Maltose-binding protein, TrxA, DsbA, hemagglutinin epitope, InaD, NorpA, and GFP, in combination with a binding partner or candidate binding partner that is tagged with at least one of the following tags: calmodulin binding peptide, streptavidin binding peptide, calmodulin binding peptide,
30 streptavidin, avidin, polyhistidine tag, polyarginine tag, FLAG tag, c-myc tag, S-tag, cellulose

binding domain, chitin-binding domain, glutathione S-transferase tag, Maltose-binding protein, TrxA, DsbA, hemagglutinin epitope, InaD, NorpA, and GFP.

The affinity tags may be fused in-frame to a protein of interest such that the tags are directly adjacent to each other, and/or to the protein of interest, or they may be separated from each other and/or from the protein of interest, by a linker (for example of 2-20 amino acids). The order in which the tags are fused with the polypeptide is not critical but can be chosen according to the affinity protocol to be used. Preferably, the tags are located near to the same end of the polypeptide(s). The location of the tag(s) is selected to allow for expression of an appropriate concentration of a correctly folded and processed tagged protein of interest. The tagged protein must not interfere with protein function, cell growth or cell viability.

Small peptides such as CBP or SBP can even be fused to the polypeptide(s) of interest internally (as long as the reading frame of the nucleic acid encoding either the tag or the nucleic acid of interest is maintained).

In one embodiment, at least one affinity tag, for example SBP is fused to a first protein and at least one affinity tag, for example CBP is fused to a second protein of the same complex. This strategy allows the purification of protein complexes containing two given proteins even when only a small fraction of the target proteins are associated, e.g., when large fractions remain free or bound to other complexes.

The invention provides for a method of detecting a binding partner ("target") for a protein of interest ("bait"). According to the method of the invention, a "bait" protein that comprises at least two tags is expressed in a cell with one or more "target" binding partners that comprise at least one different tag. In one embodiment, the bait comprises tandem, adjacent SBP and CBP tags and the binding partner comprises a third tag, for example a FLAG tag. The invention also provides for a binding partner that expresses at least one of any of the following tags: biotin, calmodulin binding peptide, streptavidin binding peptide, calmodulin binding peptide, streptavidin, avidin, polyhistidine tag, polyarginine tag, FLAG tag, c-myc tag, S-tag, cellulose binding domain, chitin-binding domain, glutathione S-transferase tag, Maltose-binding protein, TrxA, DsbA, hemagglutinin epitope, InaD, NorpA, and GFP.

Vectors

The invention provides for polynucleotides that can be provided in vectors and used for production of a tagged protein of interest. The tagged protein of interest is used, according to the methods of the invention, to purify a protein complex of interest, and/or to identify binding partners for the protein of interest.

A vector of the invention is designed to maintain expression of the chimeric protein and or candidate binding partner, at, or close to, its natural level. Overexpression of the protein may induce association with nonnatural binding partners. Transcriptional control sequences are therefore selected so that the chimeric protein is not over-expressed but is expressed at basal levels in the cell. For example, a protein of interest is expressed under the control of the endogenous promoter for the protein of interest. This serves to ensure that the protein is expressed in a native form. As used herein, “native form” means that a correct or relatively close to natural three-dimensional structure of the protein is achieved, i.e., the protein is folded correctly. More preferably, the protein will also be processed correctly and correctly modified at both the post-transcriptional and post-translational level. The correct folding is of great importance especially when the expressed polypeptide is a subunit of a protein complex because it will only bind to the other subunits of the complex when it is present in its native conformation. It is also possible to express mutant proteins, according to the methods of the invention. These can also have a native conformation. Such mutant proteins can, for example, be used to purify mutant complexes, i.e., complexes that contain some other mutated protein.

A vector of the invention contains a nucleic acid of interest under the control of sequences which facilitate the expression of the chimeric protein in a particular host cell or cell-free system. The control sequences comprise sequences such as a promoter, and, if necessary enhancers, poly A sites, etc... The promoter and other control sequences are selected so that the chimeric protein is preferably expressed at a basal level so that it is produced in soluble form and not as insoluble material. Preferably, the chimeric protein is also expressed in such a way as to allow correct folding for the protein to be in a native conformation. Preferably, one or more selectable markers are also present on the vector for the maintenance in prokaryotic or eukaryotic cells. Basic cloning vectors are described in Sambrook et al., Molecular Cloning, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press, (1989). Examples of

vectors useful according to the invention include plasmids, bacteriophages, other viral vectors and the like. Vectors useful according to the invention are also presented in Figures 3 and 4.

In a preferred embodiment, vectors are constructed containing pre-made cassettes of an affinity tag or affinity tag combinations (for example, two or more adjacent tags, wherein a first tag is an SBP tag, for example, having the nucleotide sequence presented in Figure 1, or two or more adjacent tags, wherein a first tag is an SBP tag and a second tag is a CBP tag) into which the nucleic acid coding the protein of interest can be inserted by means of a multiple cloning site such as a polynucleotide linker. Thus, a vector according to the invention is also one which does not contain the coding sequences for the protein of interest but contains the above-recited vector components plus one or more polynucleotide linkers with preferably unique restriction sites in such a way that the insertion of nucleic acid sequences, according to conventional cloning methods, into one of the sites in the polynucleotide linker, leads to a vector encoding the chimeric protein of the invention. Unique restriction enzyme sites located upstream and downstream of the tag or tags of the invention, facilitate cloning of a target protein of interest such that the tag or tags are located N- or C-terminally, or internally in the protein of interest.

In a further preferred embodiment, the vector comprises heterologous nucleic acid sequences in the form of two or more cassettes each comprising at least one of two different affinity tags, one of which is an SBP tag, for example, having the nucleotide sequence presented in Figure 1, and at least one polynucleotide linker for the insertion of further nucleic acids .

Alternatively, a vector of the invention comprises heterologous nucleic acid sequences in the form of two or more cassettes each comprising at least one of two different affinity tags, one of which is an SBP tag and one of which is a CBP tag. Such vectors can be used to express two subunits of a protein complex, each tagged with a different tag.

The invention provides for expression vectors that express the protein product of a gene of interest fused in frame to tandem tags. The tandem tags are fused in frame to either the N or C-terminus of the protein of interest. In one embodiment, a first tag is fused in frame to the N-terminus, and a second tag is fused in frame to the C-terminus of the protein of interest. Alternatively, one or more tags of the invention are fused internally to a protein of interest.

In a preferred embodiment, the invention provides for a CMV vector. The invention provides for regulatable expression systems that provides for expression of the chimeric protein

at a level that is, preferably, equivalent to the level of expression of the endogenous protein. In one embodiment the regulatable expression system is an ecdysone regulated expression system (Complete Control, Stratagene, No.:217468). In another embodiment, the system is regulatable due to the inclusion of aptamer sequences in the 5' untranslated region of, for example, the gene of interest (as described in Werstuck et al., 1988, Science, 282:296; Harvey et al., 2002, RNA, 8:452; Hwang et al., 1999, Proc Natl Acad Sci USA, 96:12997).

In another embodiment, the invention provides for a viral vector system to increase the transformation efficiency of mammalian cell lines.

Vectors useful according to the invention include CMV vectors wherein a CBP and a SBP tag are fused to the N or C terminus of the bait protein in each of the three possible reading frames. Vectors useful for expressing a CBP-SBP tagged protein of the invention are presented in Figure 3.

Vectors useful for expressing a FLAG tagged protein of the invention are presented in Figure 4 and are available from Stratagene.

Construction of vectors according to the invention employs conventional ligation techniques. Isolated plasmids of DNA fragments are cleaved, tailored and religated in the form desired to generate the plasmids required. If desired, analysis to confirm correct sequences in the constructed plasmids is performed in a known fashion. Suitable methods for constructing expression vectors, preparing *in vitro* transcripts, introducing DNA into host cells, and performing analyses for assessing expression and function are known to those skilled in the art.

Gene presence, amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA, dot blotting (DNA or RNA analysis), PCR, RT-PCR, Q-PCR, RNase Protection assays or in situ hybridization, using an appropriately labeled probe based on a sequence provided herein. Those skilled in the art will readily envisage how these methods may be modified, if desired. Standard DNA cloning procedures are, therefore, used to introduce the N or C terminal tandem tags in frame with the coding region of the protein of interest in an appropriate expression vector.

Cells

A vector of the invention can be introduced into an appropriate host cell. These cells can be prokaryotic or eukaryotic cells, e.g., bacterial cells, yeast cells, fungi or mammalian cells, and the vector or nucleic acid can be introduced (transformed) into these cells stably or transiently by conventional methods, protocols for which can be found in Sambrook et al. (supra).

DNA may be stably incorporated into cells or may be transiently expressed using methods known in the art (see Sambrook et al., supra). Stably transfected mammalian cells may be prepared by transfecting cells with an expression vector having a selectable marker gene, and growing the transfected cells under conditions selective for cells expressing the marker gene. To prepare transient transfectants, mammalian cells are transfected with a reporter gene of interest, to monitor transfection efficiency. In one embodiment, the bait vector is introduced via infection using a viral vector such as adenoviral vectors, AAV vectors, retroviral vectors or lentiviral vectors.

Vectors of the invention can be present extrachromosomally or integrated into the host genome, and used to produce recombinant cells or organisms such as transgenic animals.

Tagged Protein

The polynucleotides of the invention are useful for production of a tagged protein of interest. The tagged protein can be tagged at the N- or C-terminus, or a combination thereof, with one or more affinity tags as described herein. The tagged protein is used to purify a complex comprising the protein of interest and/or to identify binding partners for the protein of interest.

Complex of the Invention

The invention provides for methods of detecting and isolating a complex of the invention. A complex of the invention may comprise a complex of proteins or a complex of biomolecules, as defined herein. A complex of the invention comprises a protein of interest.

As used herein, “protein of interest” means any protein for which the nucleic acid sequence is known or available, or becomes available, such that it can be cloned into a nucleic acid vector which is suitable for expression in the appropriate host cells or cell-free expression systems. For purification of a protein complex, the nucleic acid sequence of at least one of the subunits of the protein complex must be known or available.

Proteins useful according to the invention include but are not limited to:

1) cell cycle regulatory proteins (for example cyclins, cdks, Rb, E2F, regulators of cyclins including p21,);

2) protein complexes involved in regulating intracellular transport (for example nuclear transport channels, transport into Golgi, transport into mitochondria);

3) proteins involved in the regulation of gene expression (for example transcription factors (e.g., p53, myc), transcription complexes (e.g., TATA binding protein complexes); transcriptional modulators (for example histone acetylases and histone deacetylases); components of snRNPs (involved in splice junction recognition); polyadenylation complexes; regulators of nuclear export of nucleic acids; RISC complex (components of the RNAi pathway);

4) growth factor receptors (EGFR, IGFR, FGFR);

5) regulators of the cytoskeleton (for example components of the focal adhesion complexes (paxillin, focal adhesion kinase); regulators of actin organization (racB);

6) viral proteins interacting with host proteins (for example EBNA2, EBNA1 of EBV, E1A/E1B of adenovirus, E6 and E7 of HPV);

7) proteins of pathogenic bacteria that bind to mammalian host cells; and

8) proteins in complexes that mediate cell/cell interactions (for example gap junctions (connexin).

A protein of interest useful according to the invention also includes lipoproteins, glycoproteins, phosphoproteins. Proteins or polypeptides which can be analyzed using the methods of the present invention include hormones, growth factors, neurotransmitters, enzymes, clotting factors, apolipoproteins, receptors, drugs, oncogenes, tumor antigens, tumor suppressors, structural proteins, viral antigens, parasitic antigens and bacterial antigens. Specific examples of

these compounds include proinsulin (GenBank #E00011), growth hormone, dystrophin (GenBank # NM_007124), androgen receptors, insulin-like growth factor I (GenBank #NM_00875), insulin-like growth factor II (GenBank #X07868) insulin-like growth factor binding proteins, epidermal growth factor TGF- α (GenBank #E02925), TGF- β (GenBank #AW008981), PDGF (GenBank #NM_002607), angiogenesis factors (acidic fibroblast growth factor (GenBank #E03043), basic fibroblast growth factor (GenBank #NM_002006) and angiogenin (GenBank #M11567), matrix proteins (Type IV collagen (GenBank #NM_000495), Type VII collagen (GenBank #NM_000094), laminin (GenBank # J03202), phenylalanine hydroxylase (GenBank #K03020), tyrosine hydroxylase (GenBank #X05290), oncogenes (ras (GenBank #AF 22080), fos (GenBank #k00650), myc (GenBank #J00120), erb (GenBank #X03363), src (GenBank #AH002989), sis GenBank #M84453), jun (GenBank #J04111)), E6 or E7 transforming sequence, p53 protein (GenBank #AH007667), Rb gene product (GenBank #m19701), cytokine receptor, Il-1 (GenBank #m54933), IL-6 (GenBank #e04823), IL-8 (GenBank #l19591), viral capsid protein, and proteins from viral, bacterial and parasitic organisms which can be used to induce an immunologic response, and other proteins of useful significance in the body.

The compounds which can be incorporated are only limited by the availability of the nucleic acid sequence for the protein or polypeptide to be incorporated. One skilled in the art will readily recognize that as more proteins and polypeptides become identified they can be integrated into the DNA constructs of the invention and used to transform or infect cells useful for producing an organized tissue according to the methods of the present invention. Therefore, a protein of interest includes the protein product of any open reading frame included in GenBank.

Protein Expression

Depending on the protein to be purified, the chimeric protein is expressed intracellularly or secreted into the culture medium. Alternatively, it might be targeted to other cell compartments such as the membrane. Depending on the protein, an appropriate method is used to extract the chimeric protein from the cells and/or medium. When a chimeric protein is expressed and targeted to a particular subcellular location, e.g., the membrane of cell organelles

or the cell membrane, these organelles or the cells themselves can be purified via the binding of these membrane proteins. It is also possible to purify cells or cell organelles via proteins naturally expressed on their surface which bind to the chimeric protein of the invention.

According to the invention it is also possible to use cell-free systems for the expression of the protein of interest. These must provide all the components necessary to effect expression of proteins from the nucleic acid, such as transcription factors, enzymes, ribosomes etc...*In vitro* transcription and translation systems are commercially available as kits so that it is not necessary to describe these systems in detail (e.g. rabbit reticulocyte lysate systems for translation). A cell-free or *in vitro* system should also allow the formation of complexes.

Protein Isolation

Various extraction procedures known in the art, and known to be compatible with purification of a protein of interest are used to prepare extracts from cells or organisms expressing the tagged target protein. Cell fractionation and/or tissue dissection can facilitate purification by providing a preenrichment step or can be used to assay specifically protein complex compositions in various tissues or cell compartments.

An extraction procedure that is useful according to the invention does not interfere with the interaction of the bait and the target proteins. For example, extraction is preferably performed in the absence of strong detergents and reducing agents, or any agent that may induce protein denaturation.

A protein extract is prepared from an appropriate cell type by first exposing the cell to either mechanical and/or chemical disruption. Mechanical disruption may include electric homogenizers, blenders, "Dounce" homogenizers, and sonicators. Chemical disruption of cells usually occurs with the use of detergents that solubilize cell membranes resulting in cell lysis.

Protease inhibitors and phosphatase inhibitors are routinely added to cell lysates, at concentrations well known in the art, to prevent proteolysis. Centrifugation is performed to separate soluble from insoluble protein and membranes, and both fractions are processed separately. Nucleic acid contaminants are usually removed from the soluble protein extract by first shearing the nucleic acid polymers or treating with DNase or a combination of DNase and

RNase. Protamine sulfate or polyethylene imine are added in various concentrations, known in the art, followed by centrifugation, resulting in a compact pellet of nucleic acid and protamine sulfate or polyethylene imine. This pellet is then discarded. The soluble protein extract is now ready for further processing.

5 The insoluble protein fraction described above can be solubilized with a variety of detergents, known in the art, and membrane proteins and analyzed.

Affinity Purification

10 The invention provides for a chimeric protein that comprises an affinity tag, and preferably at least two affinity tags. The presence of a second affinity tag is used to increase the purity following a second affinity chromatography step.

Methods of affinity purification useful according to the invention are well known in the art and are found on the world wide web at urich.edu/~jbell2/CHAPT3.html.

15 For purification according to the invention it is preferable to employ affinity chromatography using a matrix coated with the appropriate binding partner or “ligand” for the affinity tag used in that particular purification step.

20 A matrix material for use in affinity chromatography according to the invention has a variety of physical and chemical characteristics that give it optimal behavior. In terms of its physical properties it should have a high porosity, to allow maximum access of a wide range of macromolecules to the immobilized ligand. It should be of uniform size and rigidity to allow for good flow characteristics, and it must be mechanically and chemically stable to conditions used to immobilize the appropriate specific ligand. In terms of its chemical properties, it should have available a large number of groups that can be derivatized with the specific ligand, and it should not interact with proteins in general so that nonspecific adsorption effects are minimized.

25 A diverse variety of insoluble support materials are useful according to the invention, including but not limited to agarose derivatives, cellulose, polystyrene gels, cross-linked dextrans, polyacrylamide gels, and porous silicas, and beaded derivatives of agarose.

Methods of immobilizing a ligand of the invention onto a support matrix are provided on the world wide web at urich.edu/~jbell2/CHAPT3.html

In accordance with the preferred embodiment of the invention, to purify a complex comprising a chimeric protein with two affinity tags, two affinity purification steps are carried out. Each affinity step consists of a binding step in which the extracted protein is bound via one of its affinity tags, to a support material which is covered with the appropriate binding partner for that affinity tag. Unbound substances are removed and the protein to be purified is recovered from the support material. This can be done in at least two ways. Conventional elution techniques such as varying the pH, the salt or buffer concentrations and the like depending on the tag used, can be performed. Alternatively, the protein to be purified can be released from the support material by proteolytically cleaving off the affinity tag bound to the support. If the cleavage step is performed, the protein can be recovered in the form of a truncated chimeric protein or, if all affinity tags have been cleaved off, as the target polypeptide itself.

In one embodiment, biotin is added and competes for streptavidin binding sites occupied by SBP. EGTA is also added to complex with Ca^{2+} , thus disrupting the interaction between CBP and calmodulin. In other embodiments, other small molecules are added, and compete for binding sites on the affinity ligand, thereby dissociating bound protein complexes.

Elution conditions are preferably mild so that the interaction of the bait and the target is not disrupted. Preferably, non-physiological salt or pH conditions are avoided.

In one embodiment, non-specific binding proteins that naturally interact with calmodulin or streptavidin (for example naturally biotinylated proteins) are removed in a pre-purification step by incubation with avidin to bind biotinylated but not SBP tagged protein.

Protein Detection

Proteins associated with the tagged protein of interest are detected by a variety of methods known in the art.

Proteins are analyzed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and stained (either by Coomassie or by silver staining). Bands of interest are excised from the gel, and analyzed by mass spectrometry (for example as described in Honey et al., supra), either directly or following in-gel digestion, for example, with trypsin.

Associated proteins can also be identified by Western blot analysis or co-immunoprecipitation.

In certain embodiments, the eluate fraction from the affinity purification step(s) is concentrated, for example by TCA precipitation (Puig et al. supra) prior to analysis by SDS-PAGE.

5 Kits

The invention herein also contemplates a kit format which comprises a package unit having one or more containers of the subject vectors of the invention. The kit may also contain one or more of the following items: primers, buffers, affinity purification resins, instructions, and controls. Kits may include containers of reagents mixed together in suitable proportions for
10 performing the methods in accordance with the invention. Reagent containers preferably contain reagents in unit quantities that obviate measuring steps when performing the subject methods.

The vectors of the kit are provided in suitable packaging means, for example in a tube, either in solution in an appropriate buffer or in a lyophilized form.

15 Uses

The invention provides reagents and methods for identifying one or more protein binding partners or ligands that interact, either directly or indirectly, with a protein of interest.

The invention also provides for methods of detection and/or identification of a protein complex comprising two or more proteins or biomolecules.

20 The invention also provides a method of analyzing the structure and/or activity of a purified complex of one or more proteins or biomolecules. In particular, the method can be used to determine the approximate stoichiometry of proteins in a given complex.

The methods of the invention are also useful for purification of a protein complex, without disruption of the complex.

The open reading frames of the transcription factors MEF2a and MEF2c (Myosin Enhancing Factor) were cloned into the CMV-driven expression vectors described above, resulting in addition of CBP and SBP -tags either at the N-terminus or at the C-terminus of the tagged protein. These constructs act as the bait to co-purify interacting proteins. MEF2a and MEF2c were chosen because their interaction has previously been demonstrated to be detectable using a CBP/proteinA-based tandem affinity purification system (Cox et al., 2002, Biotechniques, 33:267-270; Cox et al., 2003, J. Biol. Chem., 278:15297-15303). Since members of the MEF2 family can dimerize with each other (forming homo- and hetero-dimers), MEF2a as well as MEF2c were inserted in mammalian expression vectors containing the FLAG-tag (for example as in Figure 4) as a fusion to either the N-terminus or the C-terminus of MEF2 and MEF2c, for immunodetection. These vectors provided the “target” protein in the purification procedure. The bait vectors containing either MEF2a or MEF2c were co-transfected with the target expression vectors (either Flag-tagged MEF2a or MEF2c) into COS-7 cells (as described below). MEF2a bait protein complexed with target MEF2c and MEF2c bait protein complexed with target MEF2a were purified using the tandem affinity purification reagents and purification procedure described below. Protein complexes were characterized by Western blotting and mass spectrometry.

EXAMPLE 2

EXPRESSION OF A TANDEMLY TAGGED PROTEIN

A tandemly tagged protein of interest was expressed as follows.

COS-7 cells were grown in DMEM media with 10% FBS and antibiotics (Pen/Strep) in T175 flasks overnight to 50-60 % confluency. Media was aspirated and 25 ml of fresh media was added before transfection. 30 µg of MEF2a-CBP-SBP and 30 ug of MEF2c-FLAG plasmids were diluted in 1.5 ml of serum-free DMEM media. 120 µl of Lipofectamine 2000 was diluted in 1.5 ml of serum-free DMEM media and incubated for 5 min at room temperature. The DNA and LF2000 solutions were combined and incubated for 20 min at room temperature. 3 mls of DNA-lipid complex was added to the cells and incubated at 37°C for 48 hr. Cells were washed three times with PBS. 5 ml of ice-cold PBS was then added to each flask, and the cells were scraped and transferred to a 15 ml conical tube. The cells were centrifuged at 1500x g for 10

minutes. The PBS was aspirated and 1ml of lysis buffer (described below) was added . Lysed cells were stored at -80°C. Cells from four to eight T175 flasks were used for each experiment.

EXAMPLE 3

PURIFICATION OF A PROTEIN COMPLEX

A protein complex comprising a tandemly tagged protein of interest and its binding partner was purified according to the following method.

All steps were performed at 4°C. Approximately 1×10^7 cells (1 x T175 flask) (prepared as described in example 2) were freeze thawed for 3 cycles in 1ml lysis buffer. The cells were centrifuged to pellet cell debris for 10 min at 16,000g. The cleared lysates from 4-8 flasks were pooled in a fresh tube. A 5µl sample was reserved and frozen for Western Blot analysis. To the remainder of the pooled lysate was added EDTA to a concentration of 2 mM, and β-mercaptoethanol to a concentration of 10 mM (4 µl of 0.5 M EDTA, and 0.7 µl of 14.4 M βME, for each 1000 µl of lysate) resulting in the lysates being contained in Streptavidin Binding Buffer.

100 µl of Streptavidin beads (50% slurry) for each 1 ml of lysate were washed in SBB to remove the ethanol storage buffer as follows. Beads for multiple 1 ml lysate preps were pooled and washed together in 1 ml of SBB. Beads were collected by centrifugation at 1500g for 5 minutes. The SBB wash supernatant was removed from the beads and the beads were resuspended a second time in 1ml of the indicated binding buffer. The beads were collected by centrifugation at 1500g for 5 minutes and resuspended in SBB (i.e., 100 µl SBB for each 100 µl aliquot of beads required).

100 µl of washed Streptavidin beads were added to 1ml of lysate. The tubes were rotated for 2 hr at 4 °C to allow proteins to bind to the beads. The beads were washed twice with SBB as described above. The tubes were rotated for 5 min at 4 °C to resuspend beads between centrifugations. After the final centrifugation step, SBB was removed from the pelleted beads.

100 µl of Streptavidin Elution Buffer (SEB) was added to the pelleted beads. The tubes were rotated for 30 min at 4°C to elute protein complex/es. The beads were pelleted by centrifugation at 1500g for 5 minutes. The supernatant containing the eluted proteins was carefully collected and transferred to a fresh tube. A 10 µl sample from the supernatant was reserved for Western Blot analysis.

2 µl of supernatant supplement (50 mM Magnesium acetate, 50 mM Imadazole, 100 mM Calcium chloride) was added per 100 µl of supernatant such that the eluted proteins were now suspended in Calmodulin Binding Buffer (CBB). An additional 900 µl of CBB was added to the eluted proteins. For each 1 ml of eluted proteins in CBB, 100 µl of Calmodulin Affinity Resin (50% slurry) was added. (Resin for multiple 1 ml preps was pooled and washed together in 1 ml of CBB. The resin was pelleted by centrifugation at 1500g for 5 minutes and resuspended to the original volume of 100 µl in CBB. 100 µl of washed Calmodulin Affinity Resin was added per 1 ml of eluted proteins). The tubes were rotated for 2 hr at 4°C to allow proteins to bind to the resin. The resin was washed twice with CBB as above. The tubes were rotated for 5 min at 4°C to resuspend the resin between centrifugations. After the last centrifugation step, the binding buffer was removed from the pelleted resin.

100 µl of Calmodulin Elution Buffer (CEB) was added to the pelleted Calmodulin Affinity Resin. The tubes were rotated for 30 min at 4°C to elute proteins. The resin was pelleted by centrifugation at 1500g for 5 minutes. The supernatant was carefully collected and transferred to a fresh tube. This supernatant contained the affinity purified protein complex/es.

The compositions of the buffers used in the examples presented herein are described below.

Lysis buffer:

10 mM Tris, pH 8.0

150 mM NaCl

0.1% Nonidet P-40

Add 10 µl of the protease inhibitor cocktail (Sigma, Cat.# p8340) and 10 µl of 100 mM PMSF per 1 ml of lysis buffer before use.

Streptavidin binding buffer (SBB) 250 ml

	10 mM Tris, pH 8.0	2.5 ml	1M Tris
	150 mM NaCl	7.5 ml	5M NaCl
	0.1% Nonidet P-40	2.5 ml	10% NP40
	2 mM EDTA	1 ml	0.5M EDTA
5	H ₂ O	to 250 ml	
	10 mM 2-mercaptoethanol (ME)	<i>Add 7 µl ME per 10 ml before use</i>	

Streptavidin elution buffer (SEB): SBB + 2 mM biotin.

25 ml

10	10 mM Tris, pH 8.0	0.25 ml	1M Tris
	150 mM NaCl	0.75 ml	5M NaCl
	0.1% Nonidet P-40	0.25 ml	10% NP40
	2 mM biotin	500 µl	0.1 M biotin
	H ₂ O	to 25 ml	
15	10 mM 2-mercaptoethanol	<i>Add 7 µl ME per 10 ml before use</i>	

Supernatant Supplement

1 ml

	50 mM Magnesium Acetate	100 µl	0.5 M Magnesium Acetate
	50 mM Imidazole	50 µl	1 M Imidazole
	100 mM Calcium chloride	100 µl	1 M Calcium chloride
20	H ₂ O	to 1 ml	

Calmodulin binding buffer (CBB) 250 ml

	10 mM Tris, pH 8.0	2.5 ml	1M Tris
	150 mM NaCl	7.5 ml	5M NaCl
25	0.1% Nonidet P-40	2.5 ml	10% NP40
	1 mM magnesium acetate	0.5 ml	0.5M MgAce
	1 mM imidazole	250 µl	1M Imidazole
	2 mM CaCl ₂	0.5 ml	1M CaCl ₂
	H ₂ O	to 250 ml	
30	10 mM 2-mercaptoethanol	<i>Add 7 µl ME per 10 ml before use</i>	

Calmodulin elution buffer (CEB) 25 ml

	10 mM Tris, pH 8.0	0.25 ml	1M Tris
	150 mM NaCl	0.75 ml	5M NaCl
35	0.1% Nonidet P-40	0.25 ml	10% NP40
	1 mM magnesium acetate	50 µl	0.5M MgAce
	1 mM imidazole	25 µl	1M Imidazole

5 mM EGTA
H₂O

250 µl
to 25 ml

0.5M EGTA

10 mM 2-mercaptoethanol

Add 7 µl ME per 10 ml before use

EXAMPLE 4

DETECTION OF A PROTEIN COMPLEX

A protein complex comprising a tandemly tagged protein of interest was detected.

Immunodetection

Figure 5 represents a Western blot of MEF2c-FLAG protein isolated according to the method of the invention, using the protocol described above. The data demonstrates that SBP/CBP-tagged MEF2a forms a complex with MEF2c-FLAG and that these proteins co-purify using the streptavidin and calmodulin affinity purification resins (lanes 4 and 7, respectively), as detected by the anti-FLAG antibody.

Affinity purified, isolated MEF2c was detected with an anti-Flag antibody hybridized to samples taken from each step of the affinity purification procedure. Cos-7 cells were co-transfected with two vector constructs. The first vector was MEF 2A with N-terminal tags Streptavidin Binding Peptide (SBP) and Calmodulin Binding Peptide (CBP). The second vector was MEF 2C with a FLAG peptide as an N-terminal tag. Cell lysates were prepared as described above. Lane 1 is 10 µl of lysate from 1x10⁷ Cos-7 cells lysed in 1ml of buffer. This lane shows the expression of the FLAG tag in the lysate. Lane 2 is 10 µl out of 100 µl of Streptavidin Beads after incubation and elution. This lane shows the material that remains on the beads after processing. Lane 3 is 10 µl of the 1000 µl of lysate after it has been incubated with the Streptavidin beads. This lane shows the material that is not bound by the beads. Lane 4 is 10 µl out of 100 µl of elution buffer used to elute proteins from the Streptavidin beads. This lane shows the MEF2a-MEF2c protein complex that is eluted from the streptavidin beads. Lane 5 is 10 µl out of 100 µl of Calmodulin beads after incubation and elution. This lane shows the proteins that remain on the beads after processing. Lane 6 is 10 µl of 1000 µl of material after incubation with Calmodulin Beads. This lane shows the proteins that are not bound by the Calmodulin beads. Lane 7 is 17 µl out of 100 µl of elution buffer used to elute the MEF2a-

MEF2c protein complex from the Calmodulin beads. This is the final affinity purified protein complex.

Detection of MEF2 and -MEF2c by staining

Figure 6 shows a 4-20% Tris-glycine acrylamide gel of affinity purified MEF2a/MEF2c, stained with Commassie Brilliant Blue. The right lane shows molecular weight markers. The lane on the left is affinity purified MEF2a-SBP/CBP and MEF2c-FLAG from 5×10^7 Cos-7 cells, co-transfected with vectors expressing these tagged proteins. Protein bands labeled "One" through "Four" were excised for mass spectroscopy analysis. Mass spectrometer data analysis identifies protein in bands "One" and "Two" as MEF 2A (MOWSE scores 56 and 85, respectively). Protein band "Three" is identified as MEF 2C (MOWSE score of 78). Protein band "Four" is identified as Actin (MOWSE score 175). MOWSE scores greater than 68 represent positive identification of the protein of interest.

OTHER EMBODIMENTS

Other embodiments will be evident to those of skill in the art. It should be understood that the foregoing detailed description is provided for clarity only and is merely exemplary. The spirit and scope of the present invention are not limited to the above examples, but are encompassed by the following claims.